(12) UK Patent Application (19) GB (11) 2 344 974 (13) A

(43) Date of A Publication 21.06.2000

(21) Application No 9827491.3

(22) Date of Filing 15.12.1998

(71) Applicant(s)
Roke Manor Research Limited
(Incorporated in the United Kingdom)
Roke Manor, ROMSEY, Hampshire, SO51 0ZN, United Kingdom

(72) Inventor(s)
Kathryn Gratland

(74) Agent and/or Address for Service
Siemens Group Services Limited
Intellectual Property Department, Siemens House, Oldbury, BRACKNELL, Berks, RG12 8FZ, United Kingdom

(51) INT CL^7
H04L 12/56

(52) UK CL (Edition R)
H4P PPS

(56) Documents Cited
EP 0744850 A2 **US 5764641 A**

(58) Field of Search
UK CL (Edition Q) **H4K KTK , H4P PPS**
INT CL^6 **H04L 12/56**
Online : EPODOC

(54) Abstract Title
Fair packet discard in networks

(57) In a method of implementing a fair packet discard in an unspecified bit rate (UBR) network, packets 30-44 to be transmitted on a link in the network are queued in a FIFO buffer 10 in their order of arrival in the buffer. Connections 20-28 are chosen for discard by making a list of connection numbers for the packets arriving in the buffer, Fig. 3, (not shown), selecting the first n connection numbers for discard in accordance with a threshold level 50-56, and updating the list. Discard is primarily achieved using early packet discard but partial packet discard may also be applied if the buffer fills to its maximum capacity 56.

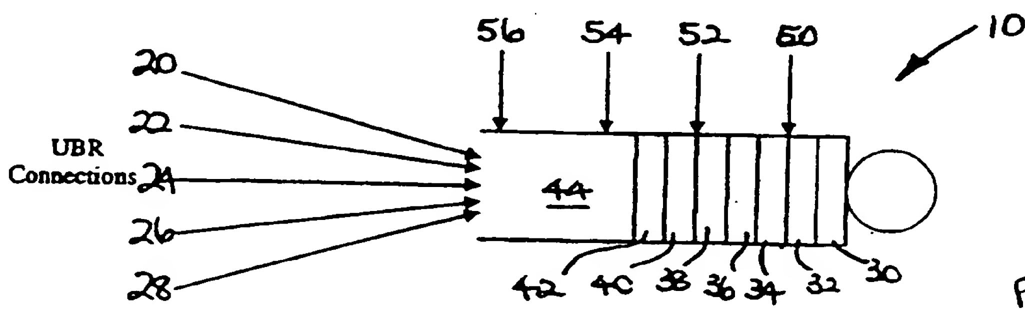

FIG. 2

GB 2 344 974 A

BNSDOCID: <GB___2344974A__I_>

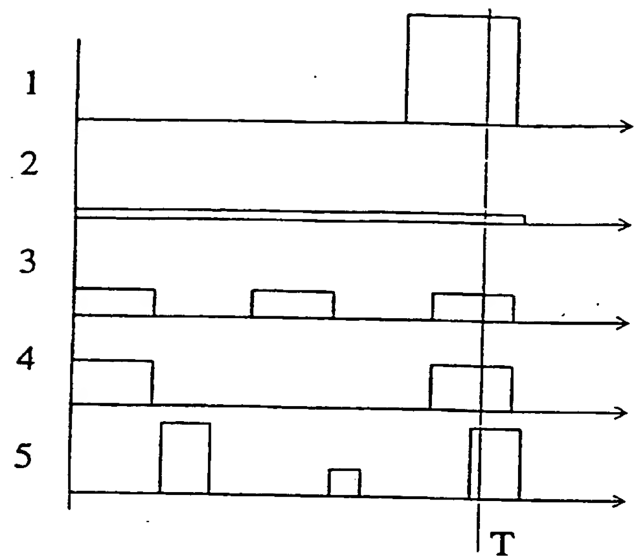

FIG. 1

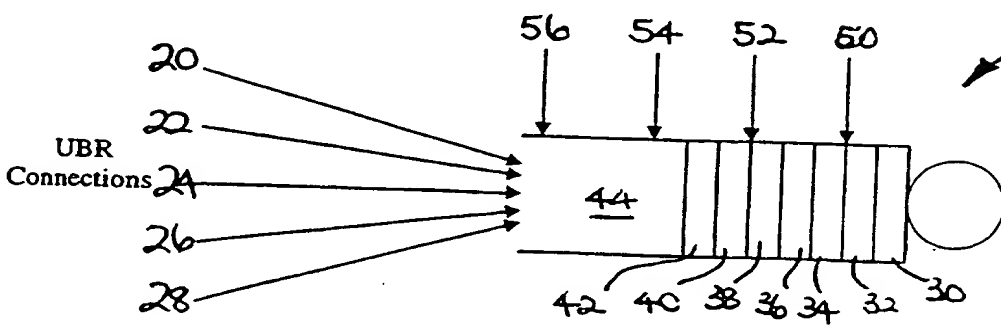

FIG. 2

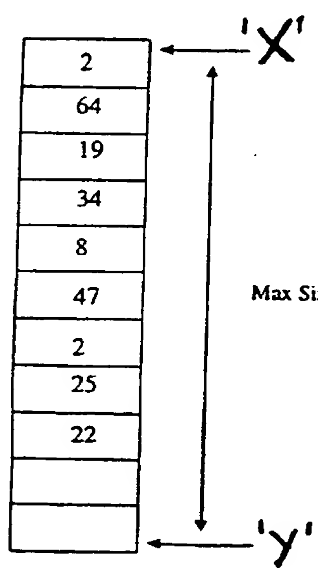

FIG. 3

IMPROVEMENTS IN OR RELATING TO NETWORKS

The present invention relates to improvements in or relating to networks.

Asynchronous transfer mode (ATM) networks typically support a mixture of all traffic types, including 'bursty' traffic. 'Bursty' traffic requires high peak bit rates for a portion of the time that the connection is active and usually little or no bandwidth for the remaining time. For example, data flow on an active connection may be on for 100 bits, off for 300 bits and on again for 100 bits giving a 25% usage. In order to efficiently allocate the network resources with respect to buffer and bandwidth in an ATM network, the 'bursty' traffic statistically shares the resources among multiple 'bursty' data flows. Due to the statistical allocation, it is possible that the instantaneous aggregate bit rate of the traffic may exceed the peak bit rate of the link. This leads to buffers filling and a congestion state occurring in the network. During a prolonged congestion state, cells may be dropped due to buffer overflow.

It is, however, preferred that the cells be discarded in accordance with a predetermined regime and there are several regimes which may be employed for discarding cells. The simplest discard regime is to 'queue tail discard'. In this regime, the network element discards single cells when the buffer becomes full. However, this typically results in poor performance for frame (or packet) based traffic streams as a single cell discard will result in the entire packet being discarded at the end of the system. If only one cell out of every ten packets is discarded (a low loss rate), this will cause the end system to discard one in ten packets which is not acceptable.

One solution to this problem is to identify each data packet and if a cell must be discarded from a packet, discard all cells from that packet thereby freeing up a large amount of buffer space. In a packet based protocol for data

BNSDOCID: <GB___2344974A__I_>

('bursty') traffic, a special 'payload type identifier' (PTI) field value is assigned to each packet. The PTI indicates which cell is the end of a packet for each connection and therefore allows the use of packet discard regimes. Packet discard makes the network utilisation more efficient as it no longer carries cells which will be discarded at the end system.

Two known regimes for discarding packet based traffic are 'early packet discard' (EPD) and 'partial packet discard' (PPD). The PPD regime maintains a threshold region in the buffer which indicates that the buffer is nearly full. This threshold region is allocated on a per connection basis or on an entire buffer basis. When the buffer occupancy exceeds the threshold, PPD drops cells from a connection. The decision on which connection to discard cells from is made either by selecting the next cell to arrive in the buffer or by using a more fair mechanism which discards based on which connection is the highest user of buffer space.

The EPD regime employs a similar mechanism to that employed by PPD, but it is more proactive. The threshold region for triggering EPD is lower than that for PPD, and when crossed, EPD drops complete incoming packets to the buffer. EPD avoids the transmission of incomplete packets and increases throughput over 'queue tail discarding'. As with PPD, the discard regime can be tailored to suit the fairness and complexity requirements of a network system. EPD may choose to discard packets from connections that are the highest users of buffer space or by selecting the connection of the next incoming cell.

In conventional implementations of EPD and PPD, a fair discard policy is based on per connection utilisation figures. In this way, an accurate choice can be made for which connections are the highest users of limited buffer space of the system. However, in a small system, for example, an access

BNSDOCID: <GB___2344974A__I_>

network, it is not always possible to provide per connection queuing or statistics on a per connection basis.

It is therefore an object of the present invention to provide an alternative mechanism for marking connections for discard in a fair manner without the large overhead required by conventional implementations of EPD and PPD.

In accordance with one aspect of the present invention, there is provided a method of implementing fair packet discard in an unspecified bit rate system, the method comprising the steps of:

a) receiving a plurality of packets into a buffer;

b) selecting one or more packets for discard when a threshold has been reached for the buffer; and

c) implementing discard for the selected packet(s);

characterised in that step a) comprises forming a queue for the received packets in order of their arrival in the buffer and step b) comprises making a list of connection numbers for the plurality of packets in the buffer, selecting the first n connection numbers for discard, and updating the list.

Advantageously, step c) comprises implementing early packet discard. Additionally, step c) may also comprise implementing partial packet discard.

For a better understanding of the present invention, reference will now be made, by way of example only, to the accompanying drawings in which:-

Figure 1 illustrates five possible unspecified bit rate (UBR) connections on a single link;

Figure 2 illustrates queuing architecture for fair packet discard (FPD) in accordance with the present invention; and

Figure 3 illustrates a discard list from which fair packet discard is implemented.

Figure 1 illustrates five possible UBR connections 1, 2, 3, 4, 5 on a single link. If, at time T, it is necessary to discard some cells because the limited resources are exhausted, a decision must be made about which connection is to have cells discarded therefrom. If there is no requirement that the discard mechanism is fair, then any connection may have the discard regime applied to it. However, if a fair discard mechanism is required, an informed choice must be made.

For the purposes of the present invention, connections with the highest utilisation will be selected for discard where several UBR connections are sharing a finite bandwidth and buffer space. However, it will be appreciated that other mechanisms can be selected to determine which connections are discarded.

Connections with the highest utilisation have been chosen for discard in this instance because:-

a) The UBR traffic does not have any weighting towards more important connections. This implies that no one connection should absorb all bandwidth and buffer space at the expense of other connections.

b) It has been shown that transmission control protocol (TCP) traffic reacts better under network congestion if a small amount of users experience high discard rather than all users experience small average discard.

In Figure 2, a shared buffer 10 comprises a single first in, first out (FIFO) device. As shown, five UBR connections 20, 22, 24, 26, 28 are directed into the buffer 10 for storage, the connections being read out in the same order that they are read in. A plurality of data packets 30, 32, 34, 36, 38, 40, 42, 44 which have already been received by the buffer 10 are also shown.

In accordance with the present invention, a method of recording the highest users on a particular link is required. A discard list as shown in Figure 3 is used to achieve this. The discard list is in the form of a ring buffer which

records the connections on which he most recent cell arrivals have occurred. As shown, the least recent arrival is indicated at 'X' and the most recent arrival at 'Y' with the maximum length of the list between these two points being 'M' entries.

When the buffer 10 fills up and an EPD threshold is crossed, the first 'n' connections from the most recent list are taken and EPD are triggered for these connections. It will be appreciated that 'n' is less than 'M' and that the discards are taken from the head of the list. For example, if 'n' is 5, connections 2, 64, 19, 34 and 8 as shown in Figure 3 will be discarded. In Figure 2, a first EPD point is indicated at 50. Similarly, two further EPD points are indicated at 52, 54. A PPD point is also indicated at 56.

As discard is performed on a per connection basis, certain parameters must be kept on a per connection basis. An example of the per connection information is shown in Table 1.

Table 1.

Connection No.	Discard Flag	EPD/PPD
1	True	Inactive
2	False	Active
3	False	Inactive
...	...	...
...	...	...
...	...	...
N	True	Inactive

It is to be noted that the method of the present invention does not require information relating to the number of cells currently in the buffer for any particular connection. The information held only indicates if there is discard active or pending for each of the connections.

For each of the 'n' connections chosen from the discard list, the 'Discard Flag' is set to 'True'. When the last cell of a packet (end of packet -

EOP) enters the buffer for this connection, the EPD is started by setting the 'EPD/PPD Flag' to 'Active' and resets the 'Discard Flag'. EPD is continued on each of these connections until the next EOP cell arrives. The 'EPD/PPD Flag' is then reset to 'Inactive'.

Although the present invention implements discard using EPD, it will be appreciated that the buffer queue may still fill up to near its maximum capacity. This is because the threshold for EPD is lower than that for PPD and although EPD may be in progress, more connections may still be sent to the buffer 10. When the threshold for PPD is reached, a reactive PPD is started. When full, the buffer must discard the next cell that enters it, and set the 'EPD/PPD Flag' to 'Active' for that connection. As with EPD, discard is 'Active' until the next EOP cell arrives in the buffer. The 'EPD/PPD Flag' is then set to 'Inactive'.

CLAIMS:

1. A method of implementing fair packet discard in an unspecified bit rate system, the method comprising the steps of:

a) receiving a plurality of packets into a buffer;

b) selecting one or more packets for discard when a threshold has been reached for the buffer; and

c) implementing discard for the selected packet(s);

characterised in that step a) comprises forming a queue for the received packets in order of their arrival in the buffer and step b) comprises making a list of connection numbers for the plurality of packets in the buffer, selecting the first n connection numbers for discard, and updating the list.

2. A method according to claim 1, wherein step c) comprises implementing early packet discard.

3. A method according to claim 1 or 2, wherein step c) comprises implementing partial packet discard.

The Patent Office

8

INVESTOR IN PEOPLE

Application No:	GB 9827491.3	**Examiner:**	Keith Williams
Claims searched:	1-3	**Date of search:**	24 May 1999

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.Q): H4P (PPS); H4K (KTK)

Int Cl (Ed.6): H04L 12/56

Other: Online EPODOC

Documents considered to be relevant:

Category	Identity of document and relevant passage		Relevant to claims
A	EP 0744850 A2	Fore Systems Inc. - see column 6, line 38 to column 7, line 43	1-3
A	US 5764641 A	Cisco Systems Inc. - see column 2, line 59 to column 4, line 22	1-3

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

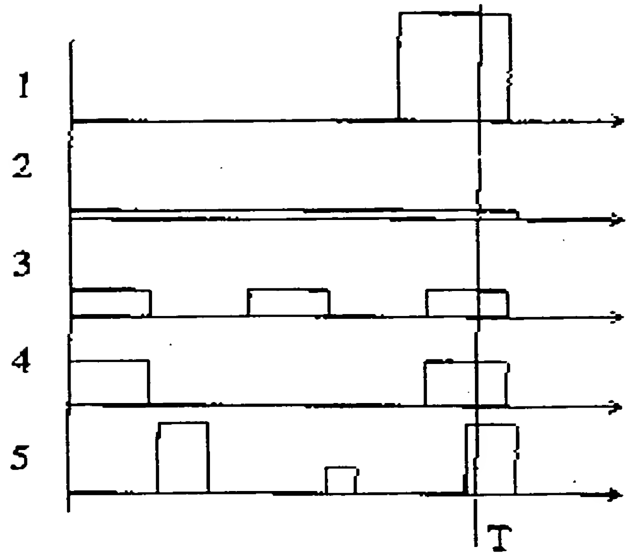

FIG. 1

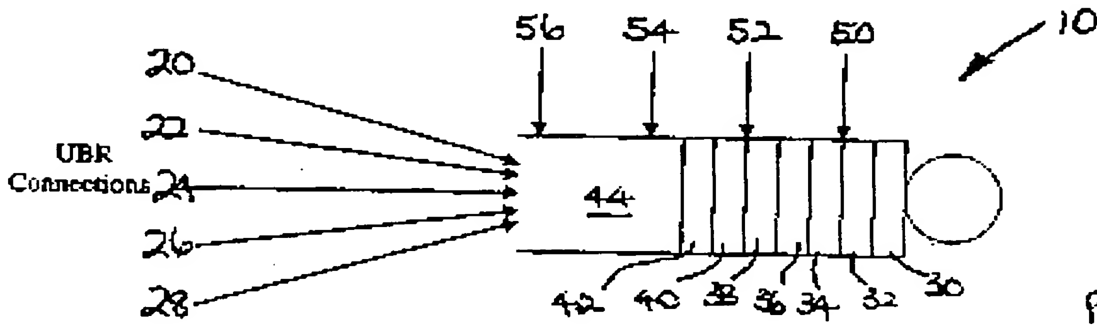

FIG. 2

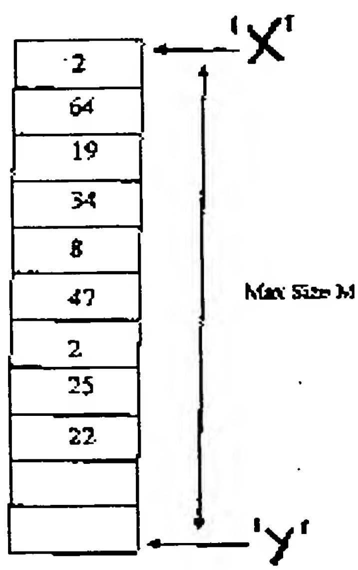

FIG. 3